JOURNAL OF MATHEMATICAL ANALYSIS AND APPLICATIONS 21, 396-404 (1968)

A Method of Samuelson and Minimal Polynomials

HSIN CHU*

Department of Mathematics, University of Maryland, College Park, Maryland 20740

Submitted by Ky Fan

1. INTRODUCTION

In this paper we study a celebrated method of Samuelson (see [1], [2], [3], or [4]) for determining explicitly the coefficients of the characteristic polynomial of a given square matrix. We find that his method is valid for a certain class of matrices only. However, it is very interesting to notice that, with some modification of his method, an effective way can be obtained for computing the coefficients of the minimal polynomial of a given square matrix as well as a given vector. To my best knowledge, as contrary to characteristic polynomials, there are no effective methods, in existing literatures available to me for determining the coefficients of a minimal polynomial, which have been discussed. A complete proof is given and examples are provided.

2. A REVIEW OF SAMUELSON'S METHOD

In the following paragraphs we give a short review of Samuelson's method.

Let $A = (a_{ij})$, $i, j = 1, 2, \ldots, n$ be a square matrix over a field F. Consider the dynamical system

$$\dot{X}(t) = AX(t), \tag{1}$$

where $X(t)$ is a column matrix with $(X(t))^t = (x_1(t), x_2(t), \ldots, x_n(t))$. Differentiating (1) $(n-1)$ times with respect to t, the following n^2 equations are obtained:

$$\begin{aligned} \dot{X}(t) &= AX(t) \\ \ddot{X}(t)\ & A\dot{X}(t) \\ &\vdots \\ X^{(n)}(t) &= AX^{(n-1)}(t). \end{aligned} \tag{2}$$

* This work was supported by Research Grant GP-6167 with the National Science Foundation.

Eliminating the $n^2 - 1$ variables not involving the subscript 1, (i.e., $x_2, \ldots, x_n$; $\dot{x}_2, \ldots, \dot{x}_n$; ...; $x_2^{(n)}, \ldots, x_n^{(n)}$) from these equations, a polynomial in x_1 and its derivatives is obtained. Substituting 1 for x_1 and $\lambda^{(i)}$ for $x_1^{(i)}(i = 1, 2, \ldots, n)$ into this polynomial, Samuelson states that the newly obtained polynomial is the characteristic polynomial of A.

3. A Condition

THEOREM 1. *A necessary and sufficient condition that Samuelson's method can be carried out and the characteristic polynomial of A can be obtained is that the following $(n-1)$ vectors in $(n-1)$-dimensional space $R, RM, \ldots, RM^{n-2}$ are linearly independent, where*

$$A = \left(\begin{array}{c|ccc} a_{11} & a_{12} & \cdots & a_{1n} \\ \hline a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdot & \cdots & \cdot \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{array}\right) = \left(\begin{array}{c|c} a_{11} & R \\ \hline S & M \\ & \end{array}\right). \tag{3}$$

PROOF: Rewriting system (2) as follows:

$$\begin{aligned}
& & -R\begin{pmatrix} x_2 \\ \vdots \\ x_n \end{pmatrix} & + (\dot{x}_1 - a_{11}x_1) & = 0 \\
& I\begin{pmatrix} \dot{x}_2 \\ \vdots \\ \dot{x}_n \end{pmatrix} & - M\begin{pmatrix} x_2 \\ \vdots \\ x_n \end{pmatrix} & - x_1 S & = 0 \\
& -R\begin{pmatrix} \dot{x}_2 \\ \vdots \\ \dot{x}_n \end{pmatrix} & + \quad 0 & + (\ddot{x}_1 - a_{11}\dot{x}_1) & = 0 \\
I\begin{pmatrix} \ddot{x}_2 \\ \vdots \\ \ddot{x}_n \end{pmatrix} & - M\begin{pmatrix} \dot{x}_2 \\ \vdots \\ \dot{x}_n \end{pmatrix} & + \quad 0 & - \dot{x}_1 S & = 0 \\
& -R\begin{pmatrix} x_2^{(n-1)} \\ \vdots \\ x_n^{(n-1)} \end{pmatrix} & + \quad 0 + \cdots & + (x_1^{(n)} - a \;\, x_1^{(n-1)}) & = 0 \\
I\begin{pmatrix} x_2^{(n)} \\ \vdots \\ x_n^{(n)} \end{pmatrix} & - M\begin{pmatrix} x_2^{(n-1)} \\ \vdots \\ x_n^{(n-1)} \end{pmatrix} & + \quad 0 + \cdots & - x_1^{(n)} S & = 0
\end{aligned} \tag{4}$$

some of the zeros here are either zero column vectors or zero row vectors. Eliminating $x_2, \ldots, x_n;\ \dot{x}_2, \ldots, \dot{x}_n; \ldots;\ x_2^{(n)}, \ldots, x_n^{(n)}$ from (4) by a standard argument in linear system, we have the following determinant:

$$\begin{vmatrix} 0 & 0 & 0 \cdots 0 & 0 & -R & (\dot{x}_1 - a_{11}x_1) \\ 0 & 0 & 0 \cdots 0 & I & -M & -x_1 S \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & -R & \cdots\ 0 & 0 & 0 & (x_1^{(n)} - a_{11}x_1^{(n-1)}) \\ I & -M & \cdots\ 0 & 0 & 0 & -x_1^{(n-1)} S \end{vmatrix} = 0$$

or

$$\begin{vmatrix} I & -M & 0 \cdots 0 & 0 & 0 & -x_1^{(n-1)} S \\ 0 & I & -M \cdots 0 & 0 & 0 & -x_1^{(n-2)} S \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 \cdots I & -M & 0 & -\dot{x}_1 S \\ 0 & 0 & 0 \cdots 0 & I & -M & -x_1 S \\ 0 & -R & 0 \cdots 0 & 0 & 0 & x_1^{(n)} - a_{11}x_1^{(n-1)} \\ 0 & 0 & -R \cdots 0 & 0 & 0 & x_1^{(n-1)} - a_{11}x_1^{(n-2)} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 \cdots 0 & -R & 0 & \ddot{x}_1 - a_{11}\dot{x}_1 \\ 0 & 0 & 0 \cdots 0 & 0 & -R & \dot{x}_1 - a_{11}x_1 \end{vmatrix} = 0. \tag{5}$$

Consequently, (5) can induce the characteristic polynomial of A if and only if the leading coefficient of $x^{(n)}$ is not zero. This means that the following determinant of order $n(n-1)$:

$$\Delta = \begin{vmatrix} I & -M & 0 \cdots 0 & 0 & 0 \\ 0 & I & -M \cdots 0 & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 \cdots I & -M & 0 \\ 0 & 0 & 0 \cdots 0 & I & -M \\ 0 & 0 & -R \cdots 0 & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 \cdots 0 & -R & 0 \\ 0 & 0 & 0 \cdots 0 & 0 & -R \end{vmatrix} \neq 0.$$

Simplifying this determinant we have

$$\Delta = \begin{vmatrix} -RM^{n-2} \\ -RM^{n-3} \\ \vdots \\ -R \end{vmatrix} = (-1)^{n-1} \begin{vmatrix} RM^{n-2} \\ RM^{n-3} \\ \vdots \\ R \end{vmatrix}.$$

It follows that $\Delta \neq 0$ if and only if $R, RM, \ldots, RM^{n-2}$ are linearly independent.

COROLLARY. *If the minimal polynomial of A is not equal to its characteristic polynomial, then R, RM,..., RM^{n-2} are linearly dependent.*

PROOF: From (3) we have

$$A = \left(\frac{a_1}{S_1}\middle|\frac{R}{M}\right),$$

where $a_1 = a_{11}$ and $S_1 = S$.
By multiplication of A repeatedly by itself we have

$$A^2 = \left(\frac{a_2}{S_2}\middle|\frac{a_1R + RM}{*}\right)$$

$$A^3 = \left(\frac{a_3}{S_3}\middle|\frac{a_2R + a_1RM + RM^2}{}\right)$$

$$\vdots$$

$$A^p = \left(\frac{a_p}{S_p}\middle|\frac{a_{(p-1)}R + a_{(p-2)}RM + \cdots + a_1RM^{p-2} + RM^{p-1}}{*}\right).$$

If the minimal polynomial of A is not equal to its characteristic polynomial, then there exists a positive integer $p < n$, and a set of numbers, α_1, α_2,..., α_p, not all zero, in the field F such that

$$A^p + \alpha_1A^{p-1} + \cdots + \alpha_pI = 0.$$

Consequently, we have

$$(a_{(p-1)}R + a_{(p-2)}RM + \cdots + a_1RM^{p-2} + RM^{p-1}) + \cdots + \alpha_{p-3}$$
$$(a_2R + a_1RM + RM^2) + \alpha_{p-2}(a_1R + RM) + \alpha_{p-1}R = 0$$

or

$$RM^{p-1} + (a_1 + \alpha_1)\, RM^{p-2} + \cdots + (a_{p-2}) + \cdots + \alpha_{p-3}a_1 + \alpha_{p-2})\, RM$$
$$+ (a_{(p-1)} + \cdots + \alpha_{p-3}a_2 + \alpha_{p-2}a_1 + \alpha_{p-1})\, R = 0.$$

It follows that RM^{p-1}, RM^{p-2},..., RM, R are linearly dependent.

The inverse statement of the corollary, in general, is not true. This can be easily seen from the following example:

EXAMPLE 1. Let

$$A = \begin{pmatrix} 2 & -1 & 1 \\ 0 & 1 & 1 \\ -1 & 1 & 1 \end{pmatrix}.$$

Its characteristic polynomial and minimal polynomial are the same and are equal to $\lambda^3 - 4\lambda^2 + 5\lambda - 2$. Where $R = (-1, 1)$ and $M = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$, we have $RM = (0, 0)$. Hence R and RM are linearly dependent.

4. The Modified Method for Finding the Minimal Polynomial of a Square Matrix

Let $A = (a_{ij})$, where $i, j = 1, 2,..., n$ be a given square matrix over a field. Consider

$$\dot{X}(t) = AX(t), \tag{6}$$

where $X(t)$ is a column vector of n components with $X^t = (x_1(t),..., x_n(t))$ and $x_i(t)$, $i = 1, 2,..., n$, are differentiable functions of t. Differentiating (6) repeatedly and substituting (6) into the resulting equations, we obtain n^2 equations

$$\begin{aligned} \dot{X}(t) &= AX(t) \\ \ddot{X}(t) &= A^2X(t) \\ &\vdots \\ X^{(n)}(t) &= A^nX(t). \end{aligned} \tag{7}$$

Theorem 2. *There exist a smallest positive integer p, $p \leqslant n$,* and *P numbers $\alpha_1, \alpha_2,..., \alpha_p$ such that $X^{(p)}(t) + \alpha_1X^{(p-1)}(t) + \cdots + \alpha_{p-1}\dot{X}(t) + \alpha_pX(t) = 0$ if and only if $\lambda^p + \alpha_1\lambda^{p-1} + \cdots + \alpha_{p-1}\lambda + \alpha_p = 0$* is the *minimal polynomial of A.*

Proof: By the Hamilton-Cayley theorem, we know

$$A^n + \beta_1A^{n-1} + \cdots + \beta_{n-1}A + \beta_nE = 0 \text{ where } \beta_1,..., \beta_n$$

are the coefficients of the characteristic polynomial of A. Consequently,

$$X^{(n)}(t) + \beta_1X^{(n-1)}(t) + \cdots + \beta_{n-1}\dot{X}(t) = (A^n + \beta_1A^{n-1} + \cdots + \beta_{n-1}A)\,X(t)$$
$$= -\beta_nX(t) \text{ or } X^{(n)}(t) + \beta_1X^{(n-1)}(t) + \cdots + \beta_{n-1}\dot{X}(t) + \beta_nX(t) = 0.$$

Let $F = \{(0,..., 0, 1, \gamma_1,..., \gamma_q) \mid X^{(q)}(t) + \gamma_1X^{(q-1)}(t) + \cdots + \gamma_qX(t) = 0, q < n.\}$ Then F is not empty. Let $(0,..., 0, 1, \alpha_1,..., \alpha_p)$, be in F where p is the smallest positive integer. It is not hard to see that this element is unique in F. From $X^{(p)}(t) + \alpha_1X^{(p-1)}(t) + \cdots + \alpha_pX(t) = 0$, we have

$$(A^p + \alpha_1A^{p-1} + \cdots + \alpha_{p-1}A + \alpha_pI)\,X(t) = 0 \text{ for all } X(t).$$

It follows that $A^p + \alpha_1A^{p-1} + \cdots + \alpha_{p-1}A + \alpha_pI = 0$. Since p is the smallest positive integer, $\lambda^p + \alpha_1\lambda^{p-1} + \cdots + \alpha_{p-1}\lambda + \alpha_p = 0$ must be minimal polynomial of A. The inverse statement is a direct consequence of the definition of minimal polynomial. The Theorem 2 is proved.

Corollary. *Let* $A = (a_{ij}^{(1)})$, $A^2 = (a_{ij}^{(2)}),..., A^n = (a_{ij}^{(n)})$. Eliminate $x_1,..., x_{i-1}, x_{i+1},..., x_n$ *from the following systems of equations*:

$$\begin{aligned} \dot{x}_i &= a_{i1}^{(1)}x_1 + a_{i2}^{(1)}x_2 + \cdots + a_{in}^{(1)}x_n \\ \ddot{x}_i &= a_{i1}^{(2)}x_1 \quad a_{i2}^{(2)}x_2 + \cdots + a_{in}^{(2)}x_n \\ &\vdots \\ x_i^{(n)} &= a_{i1}^{(n)}x_1 + a_{i2}^{(n)}x_2 + \cdots + a_{in}^{(n)}x_n, \end{aligned} \tag{8}$$

where $i = 1, 2, \ldots, n$. *Substituting* λ^j *for* $x_i^{(j)}(j = 1, 2, \ldots,)$ *and* 1 *for* x_i *we find* n *polynomials*, $\lambda^{p_i} + \alpha_{i1}\lambda^{p_i-1} + \cdots + \alpha_{ip_i}$, *where* $i = 1, 2, \ldots, n$, $p_i \leq n$. *Then the least common multiple of these* n *polynomials is the minimal polynomial of* A.

5. Some Remarks

In the actual computation, however, the numerical calculation could be considerably simpler, for examples:

(a) If $p_j = n$ for some j in (4), then $\lambda^{p_j} + \alpha_{j1}\lambda^{p_j-1} + \cdots + \alpha_{jp_j}$ is the minimal polynomial of A. That means if for some j,

$$\begin{vmatrix} a_{j1}^{(1)} & \cdots a_{jj-1}^{(1)} & a_{jj1}^{(1)} & \cdots a_{jn}^{(1)} \\ a_{j1}^{(2)} & \cdots a_{jj-1}^{(2)} & a_{jj1}^{(2)} & \cdots a_{jn}^{(2)} \\ \cdot & \cdot & \cdot & \cdot \\ a_{j1}^{(n-1)} & \cdots a_{jj-1}^{(n-1)} & a_{jj1}^{(n-1)} & \cdots a_{jn}^{(n-1)} \end{vmatrix} \neq 0,$$

then the polynomial obtained by eliminating $x_1, x_2, \ldots, x_{j-1}, x_{j+1}, \ldots, x_n$ from

$$\begin{aligned} \dot{x}_j &= a_{j1}^{(1)}x_1 + a_{j2}^{(1)}x_2 + \cdots + a_{jn}^{(1)}x_n \\ \ddot{x}_j &= a_{j1}^{(2)}x_1 + a_{j2}^{(2)}x_2 + \cdots + a_{jn}^{(2)}x_n \\ &\vdots \\ x_j^{(n)} &= a_{j1}^{(n)}x_1 + a_{j2}^{(n)}x_2 + \cdots + a_{jn}^{(n)}x_n \end{aligned}$$

is the minimal polynomial of A.

(b) If $A = \begin{pmatrix} A_1 & 0 \\ 0 & A_2 \end{pmatrix}$, find the minimal polynomials of A_1 and A_2 respectively and the least common multiple of these two polynomials is the minimal polynomial of A.

6. Examples

Example 2. Let

$$A = \begin{pmatrix} 3 & 2 & 0 & 0 \\ -4 & -1 & 0 & 0 \\ 6 & 1 & 2 & 1 \\ -1 & -3 & 1 & 0 \end{pmatrix}$$

then

$$A^2 = \begin{pmatrix} 1 & 4 & 0 & 0 \\ -8 & -7 & 0 & 0 \\ 25 & 10 & 5 & 2 \\ 15 & 2 & 2 & 1 \end{pmatrix} \qquad A^3 = \begin{pmatrix} -13 & -2 & 0 & 0 \\ 4 & -9 & 0 & 0 \\ 63 & 39 & 12 & 5 \\ 48 & 27 & 5 & 2 \end{pmatrix}$$

and

$$A^4 = \begin{pmatrix} 31 & -24 & 0 & 0 \\ 48 & 17 & 0 & 0 \\ 100 & 84 & 29 & 12 \\ 64 & 68 & 12 & 5 \end{pmatrix}.$$

Eliminate x_1, x_2, x_4 from

$$\begin{aligned} \dot{x}_3 &= 6x_1 + x_2 + 2x_3 + x_4 \\ \ddot{x}_3 &= 25x_1 + 10x_2 + 5x_3 + 2x_4 \\ \dddot{x}_3 &= 63x_1 + 39x_2 + 12x_3 + 5x_4 \\ \ddddot{x}_3 &= 100x_1 + 84x_2 + 29x_3 + 12x_4 . \end{aligned}$$

By the well-known Crout's formulation we have

$$\left[\begin{array}{ccc|ccccc} 6 & 1 & 1 & 0 & 0 & 0 & -1 & 2 \\ 25 & 10 & 2 & 0 & 0 & -1 & 0 & 5 \\ 63 & 39 & 5 & 0 & -1 & 0 & 0 & 12 \\ 100 & 84 & 12 & -1 & 0 & 0 & 0 & 29 \end{array}\right]$$

$$\begin{array}{c|ccc|ccccc|} 6 & 1 & \frac{1}{6} & \frac{1}{6} & 0 & 0 & 0 & -\frac{1}{6} & \frac{2}{6} \\ & 25 & 10 & 2 & 0 & 0 & -1 & 0 & 5 \\ & 63 & 39 & 5 & 0 & -1 & 0 & 0 & 12 \\ & 100 & 84 & 12 & -1 & 0 & 0 & 0 & 29 \\ \hline \frac{35}{6} & & 1 & -\frac{13}{35} & 0 & 0 & -\frac{6}{35} & \frac{25}{35} & -\frac{20}{35} \\ & & \frac{57}{2} & -\frac{11}{2} & 0 & -1 & 0 & \frac{21}{2} & -\frac{18}{2} \\ & & \frac{202}{3} & -\frac{14}{3} & -1 & 0 & 0 & \frac{50}{3} & -\frac{13}{3} \\ \hline \frac{356}{70} & & & 1 & 0 & -\frac{70}{356} & \frac{342}{356} & -\frac{690}{356} & \frac{510}{356} \\ & & & \frac{2136}{105} & -1 & 0 & \frac{1212}{105} & -\frac{3300}{105} & \frac{3585}{105} \\ \hline -1 & & & & 1 & -4 & 8 & -8 & -5 \end{array}$$

The minimal polynomial of A is $\lambda^4 - 4\lambda^3 + 8\lambda^2 - 8\lambda - 5$.

EXAMPLE 3. Let

$$A = \begin{pmatrix} 3 & -3 & 2 \\ -1 & 5 & -2 \\ -1 & 3 & 0 \end{pmatrix}.$$

Then

$$A^2 = \begin{pmatrix} 10 & -18 & 12 \\ -6 & 22 & -12 \\ -6 & 18 & -8 \end{pmatrix} \quad \text{and} \quad A^3 = \begin{pmatrix} 36 & 84 & 56 \\ -28 & 92 & -56 \\ -28 & 84 & -48 \end{pmatrix}$$

Eliminating x_2, x_3 from the following equations:

$$\begin{aligned} \dot{x}_1 &= 3x_1 - 3x_2 + 2x_3 \\ \ddot{x}_1 &= 10x_1 - 18x_2 + 12x_3 \\ \dddot{x}_1 &= 36x_1 - 84x_2 + 56x_3, \end{aligned}$$

we have $\ddot{x}_1 - 6\dot{x}_1 + 8x_1 = 0$. Its minimal polynomial is

$$\lambda^2 - 6\lambda + 8 = 0.$$

7. THE MODIFIED METHOD FOR FINDING THE MINIMAL POLYNOMIAL OF A VECTOR

Similarly, we can modify the Samuelson's method for finding the minimal polynomial of an n-dimensional vector over a given field with respect to a linear operator A. Since the proof is similar to Theorem 2, we state the result without proof.

Let $A = (a_{ij})$ with $i, j = 1, 2, \ldots, n$ be a given square matrix over a field. Let b be a given vector with $b^t = (b_1, \ldots, b_n)$. Consider

$$\dot{X}(t) = A^t X(t), \tag{9}$$

where $X(t)$ is a column vector of n components with $X^t = (x_1(t), \ldots, x_n(t))$ and $x_i(t) (i = 1, 2, \ldots, n)$ are n-times differentiable functions of t.

Let $y = b^t X$.

Then, by (9), we have

$$\left\{ \begin{aligned} \dot{y} &= b^t A^t X \\ \ddot{y} &= b^t (A^t)^2 X \\ &\vdots \\ y^{(n)} &= b^t (A^t)^n X \end{aligned} \right\} \tag{10}$$

and

THEOREM 3. *There exist a smallest positive integer* p, $p \leqslant n$, and p *numbers* $\alpha_1, \alpha_2, \ldots, \alpha_p$ *such that* $y^{(n)} + \alpha_1 y^{(n-1)} + \cdots + \alpha_{p-1}\dot{y} + \alpha_p y = 0$ *if and only if* $\lambda^p + \alpha_1\lambda^{p-1} + \cdots + \alpha_{p-1}\lambda + \alpha_p = 0$ *is the minimal polynomial of* b *with respect to* A.

COROLLARY. *Eliminate* $x_1, x_2, \ldots, x_n$ *from* (10):

$$\left\{\begin{aligned} y &= b^t X \\ \dot{y} &= b^t A^t X \\ \ddot{y} &= b^t (A^t)^2 X \\ &\vdots \\ y^{(n)} &= b^t (A^t)^n X \end{aligned}\right\}.$$

Substituting λ^j *for* $y^{(j)}$, $j = 1, 2, \ldots, n$, *and* 1 *for* y, *we find the minimal polynomial* $\lambda^p + \alpha_1\lambda^{p-1} + \cdots + \alpha_{p-1}\lambda + \alpha_p = 0$ *of* b *with respect to* A.

REFERENCES

1. E. BODEWIG. "Matrix Calculus." Wiley (Interscience), New York, 1959.
2. A. S. HOUSEHOLDER. "The Theory of Matrices in Numerical Analysis," Blaisdell, 1964.
3. P. A. SAMUELSON. A method of determining explicitly the coefficients of the characteristic equation. *Ann. Math. Stat.* 13 (1942), 424-429.
4. H. WAYLAND. Expansion of determinantal equations into polynomial Form. *Quart. Appl. Math.* II (1945), 277-306.
5. R. BELLMAN. "Introduction to Matrix Analysis." McGraw-Hill, New York, 1960.
6. F. R. GANTMACHER. "Matrix Theory," Vols. I and II. Chelsea, New York, 1959.
7. A. S. HOUSEHOLDER AND F. L. BAUER. On certain methods for expanding the characteristic polynomial. *Numer. Math.* 1 (1959), 29-37.